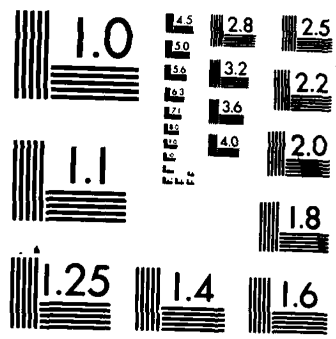
AD-A160 344 DISCREPANCY-TOLERANT HIERARCHICAL POISSON EVENT-RATE ANALYSES(U) NAVAL POSTGRADUATE SCHOOL MONTEREY CA D P GAVER JUL 85 NPS55-85-016 1/1

UNCLASSIFIED F/G 12/1 NL

END

FILMED

DTIC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

NPS55-85-016

NAVAL POSTGRADUATE SCHOOL

Monterey, California

DISCREPANCY-TOLERANT HIERARCHICAL POISSON
EVENT-RATE ANALYSES

by

Donald P. Gaver

July 1985

Prepared for:
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217-5000

NAVAL POSTGRADUATE SCHOOL
Monterey, California

Reproduction of all or part of this report is authorized.

Donald P. Gaver

DONALD P. GAVER, Professor
Dept. of Operations Research

Reviewed by:

Alan R. Washburn

Alan R. Washburn, Chairman
Department of Operations Research

Released by:

Kneale T. Marshall

Kneale T. Marshall
Dean of Information and Policy
Sciences

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: NPS55-85-016
2. GOVT ACCESSION NO.: AD-A160 344
3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle): DISCREPANCY-TOLERANT HIERARCHICAL POISSON EVENT-RATE ANALYSES
5. TYPE OF REPORT & PERIOD COVERED: Technical
6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s): Donald P. Gaver
8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Naval Postgraduate School, Monterey, CA 93943-5100
10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: 61153N; RR014-05-01 N0001485WR24061
11. CONTROLLING OFFICE NAME AND ADDRESS: Office of Naval Research, 800 North Quincy Street, Arlington, VA 22217-5000
12. REPORT DATE: July 1985
13. NUMBER OF PAGES: 49
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)
15. SECURITY CLASS. (of this report): Unclassified
15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report): Approved for public release; distribution unlimited.
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)
18. SUPPLEMENTARY NOTES
19. KEY WORDS (Continue on reverse side if necessary and identify by block number): Statistical estimation, Poisson process, empirical Bayes
20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

There are $J>1$ units (machines) that generate events (failures) at possibly different, constant, Poisson rates. Having observed a record of such events, it is desired to (a) characterize the overall variability of "true rates", and (b) use the result of (a) to create improved estimates of the individual rates by selective pooling. The results are evaluated by simulation, and applied to actual operational data.

DD FORM 1473, 1 JAN 73 EDITION OF 1 NOV 65 IS OBSOLETE

S/N 0102-LF-014-6601

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

DISCREPANCY-TOLERANT HIERARCHICAL POISSON EVENT-RATE ANALYSES

Donald P. Gaver

1. Introduction

Consider a collection of J entities called "units" that independently generate events in accordance with Poisson processes, each with rate parameter λ_j. We are in possession of observations on each of these processes, and have seen s_j events for a time exposure of t_j for the $j^{\underline{th}}$, $j = 1,\ldots,J$. Possibly also available are concommitant observations on other variables $\underline{x}$ that may in part influence (explain) the values of the rates λ_j. The problem is to use these observations to describe the nature and extent of the variation between individual unit rates, and on this basis to predict (a) the future event generation behavior of individual units under observation, as well as (b) the overall rate variability of existing units, and hence the likely rate behavior of other, similar, units not yet under observation. The object of this paper is to propose and examine statistical methods for approaching the above problems. The approach emphasized is to treat the unknown rates as being describable in part as coming from a fixed population of possible rates, and then to describe or assess that population and its implications for estimating the individualized unit rates. The approach is called <u>hierarchical</u> because each rate may be viewed as a realization of some random variable associated with a higher-level <u>superpopulation</u> of rates; such models are also called <u>random parameter</u> or <u>parametric empirical Bayes</u> models; see Morris (1983)

for a review. There have been a variety of applications of similar models in many fields. However, particular emphasis is given in this paper to analyses that invoke <u>discrepancy</u> <u>tolerant</u> superpopulation parametric representations: ones yielding estimating procedures that may assist in identification of distinct rate groupings, existence of apparently discrepant or outlying rates, etc., a better understanding of which could suggest desirable improvements for systems so identified. Of course this latter step may well lead in practice to a change in the superpopulation, and to need for an updated new analysis. The steps suggested resemble the cycle of data analysis and modelling, model diagnosis by residual and sensitivity analysis, and repeat, often adopted in enlightened regression analyses; cf. Mosteller and Tukey (1977), Belsley, Kuh, and Welsch (1980), and elsewhere. Some ideas of discrepancy-tolerant or robust Bayesian analyses have been described by Berger (1980), (1984), who references Albert (1979) for as-yet unpublished studies. Ideas expressed in the paper of Box (1980), with discussion, are quite relevant, as well.

This paper proceeds by first introducing hierarchical Poisson models. Specification of useful parametric forms for the superpopulation that describes between-unit variability is the next topic; this is followed by a discussion of explicit adjusted estimates for individual event rates in terms of superpopulation parameters. Finally, some procedures are described for obtaining estimates of the superpopulation parameters. The estimation procedure effectiveness is assessed by simulation, and the technology is applied to certain sets of real data.

2. The Hierarchical Poisson Model

Introduce as a starting point the Poisson process of events for item j with (conditional) event rate $\lambda_j(\underline{x}_j, \varepsilon_j)$ where $\underline{x}_j = (x_{1j}, x_{2j}, \ldots, x_{pj})$ is a vector of explanatory (regression) variables, and ε_j is (the realization of) a random variable with fixed density $f_{\varepsilon_j}(x;\underline{\theta})$; the latter describes an infinite superpopulation with the parameters vector $\underline{\theta}$. The value of the $j^{\underline{th}}$ $(1 \leq j \leq J)$ latent variable ε_j is here taken to be fixed for all time, once drawn from the superpopulation. It is thus a <u>random individualization</u> of the failure rate of item j; while $\underline{x}_j$ values account for differences in environmental factors. Generally, the superpopulation distribution accounts for the rate variations between items or individuals after adjustment for environmental effects explained by $\underline{x}_j$. Of course the manner in which the explanatory variables are used can influence the form of the apparent superpopulation distribution, and the selection of items for study, if influenced by the rate values, can bias the estimation of superpopulation parameters; see Lehoczky (1984).

Contrast the above model types to those in which λ_j is a random function of time: e.g. in discrete time, monthly or weekly perhaps, the rate changes in accordance with $\lambda_{jt} \equiv \lambda_j(\underline{x}_j, \varepsilon_{jt})$, and $\{\varepsilon_{jt}, t = 1,2,\ldots\}$ is a collection of possibly iid random variables, or perhaps a more general stochastic process changing in discrete or continuous time; these last are called "random environment" models, or more specifically doubly stochastic Poisson models, see Cox and Lewis (1966), Cox and Isham (1980), Gaver (1963), Reynolds and Savage (1971), and Burridge (1981). For instance, if the

integrated hazard $\int_0^t \lambda'_{jt}\,dt' = \Lambda_j(t)$ is the realization of a gamma stochastic process, then the original Poisson process becomes a negative binomial process. Other interesting models would allow for changes in the superpopulation as a result of event observation and remedial action. Consideration of all of these latter is beyond the scope of this study; this paper confines its attention to the simplest random individualization hierarchical structures.

3. Two Classical Hierarchical Models: The Log-Normal and Gamma Super-populations

In order to parameterize a superpopulation of even rates in a simple fashion, both the log-normal and the gamma distribution have been utilized historically. Here is a brief discussion, with modifications reserved for the following section.

3.1 Log-Normal Superpopulation: The L/N/P (Log-Normal Poisson) Model

Suppose that the Poisson rate of item j is of the form

$$\tilde{\lambda}_j = \exp(\tilde{\varepsilon}_j) \qquad j = 1,2,\ldots,J \tag{3.1}$$

where $\tilde{\varepsilon}_j \sim N(\mu_j,\sigma^2)$. Let $\mu_j = \mu + \underline{x}_j\underline{\beta}$, $\underline{x}_j$ being a vector of covariates and $\underline{\beta}$ a vector of regression coefficients, whenever the regression term is present; otherwise $\mu_j = \mu$. This paper does not consider the fitting of regression coefficients. Each of the J items is exposed for known time t_j, with s_j (= 0,1,2,...) observed events recorded therein.

This (log-normal) model is expecially popular in the Probabilistic Risk Assessment (PRA) of nuclear reactor safety and operating systems, see the Reactor Safety Study (WASH 1400) (1975), and subsequent numerous reports on this topic; in particular note Kaplan (1983). Items may be in-plant equipments such as continuously acting pumps, valves, and control devices that are subject to failure events; other events of concern are so-called <u>initiating events</u> such as loss of feedwater, pipe breaks, loss of offsite power and other challenges to the integrity of a nuclear--or other--plant's safe and productive operation. The failure or initiating event occurrences may initially be taken to be time-homogeneous Poisson stochastic processes, with rates that vary between design copies in accordance with environmental influences

(manufacturer, geographical location, plant type, (sub)system type, etc.). A full analysis endeavors to estimate the influence of the explanatory environmental variables as represented by a regression function, and the properties of the random individualizations, ε_j, so as to provide (i) good estimates of the superpopulation center (mean) and variability (variance) and (ii) good estimates of the individual item failure rates. An individual rate estimate constructed from its own experience alone may be subject to considerable random error; pooling its data with that of other similar items can reduce that random error at the possible expense of adding bias. Our methods suggest pooling that is reasonably based on superpopulation characteristics as estimated under (i), <u>assuming</u> that the superpopulation is (log)normal; in a later section estimates are developed that depend less on such a special form, i.e. that pool more selectively.

Given the model (3.1) a revised, pooled or shrunken, estimate of the rate associated with item j can be obtained by Bayes' formula, perhaps in the form

$$\hat{\lambda}_j = E[\underset{\sim}{\lambda}_j | s_j, \underline{x}_j] = K_j \int_{-\infty}^{\infty} \lambda(y) e^{-\frac{1}{2}\left(\frac{y-\mu_j}{\sigma}\right)^2} e^{-\lambda(y) t_j} (\lambda(y))^{s_j} dy, \tag{3.2}$$

where $\lambda(y) = \exp(y)$, $\mu_j = \mu + \underline{\beta}\, \underline{x}_j$, and K_j is a normalizing constant. The integral sometimes may be well approximated by use of Laplace's method, see Tierney and Kadane (1984). However, a likelihood approach provides quick and interpretable results: choose $\hat{\varepsilon}_j = \ell n\, \hat{\lambda}_j$ to maximize

$$L_j(\varepsilon_j;\mu,\sigma^2,s_j,t_j) = \frac{e^{-\frac{1}{2}\left(\frac{\varepsilon_j-\mu_j}{\sigma}\right)^2}}{\sqrt{2\pi\sigma}} e^{-\lambda(\varepsilon_j)t_j} \frac{(\lambda(\varepsilon_j)t_j)^{s_j}}{s_j!}, \quad (3.3)$$

the likelihood of ε_j given μ,σ^2 and the observations. Differentiate the log of (3.3) and equate to zero to obtain the non-linear estimating equation

$$e^{\hat{\varepsilon}_j} = (s_j + \frac{\mu_j-\hat{\varepsilon}_j}{\sigma^2}) \frac{1}{t_j} \quad (3.4)$$

The nature of the estimate is appreciated if we let $\hat{\varepsilon}_j(1) = \ln(s_j/t_j)$ be an initial solution (putting $s_j/t_j = 1/(3t_j)$ if $s_j = 0$), and pass through one Newton-Raphson iteration to obtain

$$\hat{\varepsilon}_j \simeq \frac{s_j \ln(s_j/t_j) + \mu_j/\sigma^2}{s_j + (1/\sigma^2)} \quad (3.5)$$

Since a delta-method approximation to $\mathrm{Var}[\ln(\underset{\sim}{s}_j/t_j) | \underset{\sim}{\lambda}_j = \lambda_j]$ is just $1/\lambda_j t_j \approx 1/s_j$, the estimate (3.5) is the linearly weighted shrinkage of the log-rate estimate towards the assumed superpopulation mean, μ, with weights the reciprocals of the within and between (superpopulation) variances, familiar in the normal/Gaussian distribution Bayesian-conjugate prior framework.

The above estimates require values for superpopulation parameters $\mu_j = \mu + \underline{x}_j\underline{\beta}$ and σ^2. Once these are at hand, values can be computed for $\hat{\varepsilon}_j$. Desirably, μ_j and σ^2 can be estimated from data on all J items' histories. Various options exist for this; some are proposed and explored in a later section.

3.2 Gamma Superpopulation

Consider the convenient (conjugate) alternative model

$$\lambda_j = U_j \exp(\beta \underline{x}_j)$$

where now all is as before except that $U_j \sim \text{Gamma}(\delta,\alpha)$, meaning that the density is, for $u \geq 0$,

$$f(u;\delta,\alpha) = e^{-\delta u} \frac{(\delta u)^{\alpha-1}}{\Gamma(\alpha)} \delta \tag{3.6}$$

This model has a long history, for it leads to the negative binomial marginal distribution of event counts:

$$P\{\underset{\sim}{s} = k \mid \underline{x}\} = E[e^{-\lambda(\underset{\sim}{U})t} \frac{[\lambda(\underset{\sim}{U})t]^k}{k!}]$$

$$= \int_0^\infty e^{-\lambda(u)t} \frac{[\lambda(u)t]^k}{k!} f(u;\delta,\alpha)du \tag{3.7}$$

$$= \frac{\Gamma(k+\alpha)}{k!\Gamma(\alpha)} \left(\frac{\delta}{\delta + t\exp(\underline{\beta}\ \underline{x})}\right)^{\alpha} \left(\frac{t\exp(\underline{\beta}\ \underline{x})}{\delta + t\exp(\underline{\beta}\ \underline{x})}\right)^{k}$$

and to a simple linear formula for the Bayes estimate of the individualized rate

$$\hat{\lambda}_j \equiv E[\underset{\sim}{\lambda}_j \mid s_j, t_j \underline{x}_j] = \frac{s_j \exp(\underline{\beta} x_j) + \alpha \exp(\underline{\beta}\ \underline{x}_j)}{t_j \exp(\underline{\beta}\ \underline{x}_j) + \delta} \tag{3.8}$$

Since for the gamma,

$$m_U \equiv E[U] = \frac{\alpha}{\delta}; \qquad \sigma_U^2 = \text{Var}[U] = \frac{\alpha}{\delta^2}, \tag{3.9}$$

then

$$\delta = \frac{m_U}{\mathrm{Var}[U]} ; \qquad \alpha = \frac{m_U^2}{\mathrm{Var}[U]} ,$$

the above expression can be expressed as

$$\hat{\lambda}_j = \frac{\left(\frac{\exp(\underline{\beta}\,\underline{x}_j)\,t_j}{m_U}\right)\cdot\left(\frac{s_j}{t_j}\right) + \frac{1}{\mathrm{Var}[U]}(m_U \exp(\underline{\beta}\,\underline{x}_j))}{\frac{\exp(\underline{\beta}\,\underline{x}_j)\,t_j}{m_U} + \frac{1}{\mathrm{Var}[U]}} \tag{3.10}$$

which is again a linear shrinking of the raw point estimate (s_j/t_j) towards the appropriate superpopulation mean, here expressed as $m_U \exp(\underline{\beta}\,\underline{x}_j)$; the weights are again recognizable as within $(m_U/\exp(\underline{\beta}\,\underline{x}_j)t_j)$ and between $(\mathrm{Var}[U])$ variance components.

Note that the random log-linear model (3.1) is only one suggestive model form. Conceivably random individualization should be of multiplicative form: $\ln \lambda_j = \varepsilon_j\,\underline{\beta}\,\underline{x}_j$, rather than additive, so that covariate influence varies from item to item. Other possibilities also exist, but regression effects are not considered further in this paper.

Option (A) (pseudo-t).

For this representation,

$$Q'(z) = \alpha\beta z\left(\frac{1+\phi^2/\alpha}{\phi}\right) \tag{5.28}$$

where $\alpha = n-2$, $\beta = (n-3/2)/(n-1)^2$ from Gaver and Kafadar (1984) requires solving the following equation for ϕ:

$$e^{\mu+\sigma\phi} = \left(s_j - \frac{(n-1)^2\phi/\sigma}{(n-2)(n-3/2)\left(1+\phi^2/(n-2)\right)}\right)\frac{1}{t_j}\ ; \tag{5.29}$$

this has arisen earlier as (4.3) in the context of finding an individualized estimate of λ_j.

To simplify calculations, one may initially take as an approximate solution

$$\hat{\phi}_j = \frac{\ln(s_j/t_j) - \mu}{\sigma}\ . \tag{5.30}$$

Reference to (4.3) then shows that

$$\hat{\theta}_j = \left(\sqrt{\frac{1}{\beta}\ln\left(1+\frac{\hat{\phi}^2}{\alpha}\right)}\right)\ \mathrm{sign}(\hat{\phi}_j) \tag{5.31}$$

Furthermore, for this determination of $\hat{\theta}_j$,

$$Q(\hat{\theta}_j) = \frac{1}{2}(\hat{\theta}_j)^2 \tag{5.32}$$

$$Q'(\hat{\theta}_j) = \hat{\theta}_j \tag{5.33}$$

evaluated by Gauss-Hermite integration. Finally, then, the log-likelihood is of the form

$$\ell(\mu,\sigma^2) = \frac{1}{J}\sum_{j=1}^{J}\{-Q(\hat{\theta}_j) + \frac{[Q'(\hat{\theta}_j)]^2}{2Q''(\hat{\theta}_j)} - \frac{1}{2}\ln Q''(\hat{\theta}_j) + \ln I_j\}, \qquad (5.26)$$

which can be examined for maxima over μ and $\sigma^2 > 0$. NOte that if the equation $Q_j'(z)$ is in fact solved precisely, then $Q'(\hat{\theta}_j) = 0$, and the correction term $[Q'(\hat{\theta}_j)]^2/2Q''(\hat{\theta}_j)$ may be omitted.

An important part of the computation involves finding $\hat{\theta}_j$, an approximate solution to the equation

$$Q_j'(z) = 0 = z + \lambda'(z)t_j - s_j\frac{\lambda'(z)}{\lambda(z)}.$$

Parametrization by the sculptured form $\ln \lambda(z) = \mu + \sigma\phi(z)$ leads to the equation

$$0 = z + \sigma(t_j\lambda(z) - s_j)\phi'(z) \qquad (5.27,a)$$

or

$$0 = z + \sigma(t_j e^{\mu+\sigma\phi(z)} - s_j)\phi'(z) \qquad (5.27,b)$$

At this point it becomes desirable to introduce a specific, tractable, parametric form for $\phi(z)$. Consult (4.3): option (A), a pseudo-t, is handy, and will be adopted. Here are some details.

regard $q_j(z)$ as the error in the quadratic approximation. Next complete the square in the quadratic terms of (5.19) and put

$$x = \frac{1}{\sqrt{2}}[(z-\hat{\theta}_j)\sqrt{Q''(\hat{\theta}_j)} + Q'(\hat{\theta}_j)/\sqrt{Q''(\hat{\theta}_j)}] \tag{5.21}$$

and hence

$$z(x) = \frac{\sqrt{2}\,x}{\sqrt{Q''(\hat{\theta}_j)}} - \frac{Q'(\hat{\theta}_j)}{Q''(\hat{\theta}_j)} + \hat{\theta}_j \tag{5.22}$$

It thus follows that

$$Q_j(z(x)) = q_j(z(x)) - \frac{[Q'(\hat{\theta}_j)]^2}{2Q''(\theta_j)} + x^2 , \tag{5.23}$$

and that the likelihood assumes the form

$$L_j(\mu,\sigma^2) = K_j \exp(-Q_j(\hat{\theta}_j) + \frac{[Q_j'(\hat{\theta}_j)]^2}{2Q_j''(\hat{\theta}_j)}) \frac{1}{\sqrt{Q_j''(\hat{\theta}_j)}} I_j(\mu,\sigma^2), \tag{5.24}$$

where

$$I_j(\mu,\sigma^2) = \int_{-\infty}^{\infty} e^{-x^2} e^{-q_j(z(x))} dx , \tag{5.25}$$

and K_j is a constant independent of parameters μ and σ^2. The idea is that subtraction of the quadratic approximation from $Q_j(z)$ should leave a relatively minor correction, I_j, to be

$$Q''_{Pj}(\theta_j) = s_j\sigma^2[\phi'(\theta_j)]^2$$

$$= s_j\sigma^2\theta_j^2\left[\frac{(n-2)(n-3/2)}{(n-1)^2}\left(\frac{1+\phi^2(\theta_j)/(n-2)}{\phi(\theta_j)}\right)\right]^2 \qquad (5.12)$$

each of which is evaluated from (5.8,b). Note that if $n \to \infty$ then the normal superpopulation case is obtained, and the present approximation treats $\log(s_j/t_j)$ as approximately $N(\mu,\sigma^2+1/s_j)$.

Method 3: Quadratic Approximation to the Log-Likelihood, Augmented by Gauss-Hermite Integration

The previous method blithely approximates the Poisson log-likelihood by a quadratic in order to achieve convenient and interpretable results. Consider next a more careful approach that combines Methods 1 and 2. To do so, express the entire exponent in the integrand of (5.4)--essentially the negative log-likelihood--as

$$Q_j(z) = \tfrac{1}{2}z^2 + Q_P(z) = \tfrac{1}{2}z^2 + t_j\lambda(z) - s_j \ln\lambda(z) . \qquad (5.18)$$

Let θ_j be a solution of $Q'_j(z) = 0$. NOte that this solution will not be as explicit as before, and so an approximate value may sometimes be most conveniently used; call it $\hat{\theta}_j$. Now expand up to quadratic terms and let

$$Q_j(z) = q_j(z) + Q(\hat{\theta}_j) + (z-\hat{\theta}_j)Q'_j(\hat{\theta}_j) + \tfrac{1}{2}(z-\hat{\theta}_j)^2Q''_j(\hat{\theta}_j) \qquad (5.19)$$

where

$$q_j(z) = Q_j(z) - Q_j(\hat{\theta}_j) - (z-\hat{\theta}_j)Q'_j(\hat{\theta}_j) - \tfrac{1}{2}(z-\hat{\theta}_j)^2Q''_j(\hat{\theta}_j) ; \qquad (5.20)$$

or

$$\phi(\theta_j) = \frac{\ln(s_j/t_j) - \mu}{\sigma} , \tag{5.8,b}$$

namely

$$\theta_j = \phi^{-1} \left(\frac{\ln(s_j/t_j) - \mu}{\sigma}\right) . \tag{5.8,c}$$

Clearly $Q_{Pj}(\theta_j)$ is independent of μ and σ^2 and can be ignored, while

$$Q''_{Pj}(\theta_j) = s_j \sigma^2 [\phi'(\theta_j)]^2 , \tag{5.9}$$

and so the approximate likelihood is of the form

$$\begin{aligned} \tilde{L}_j(\mu,\sigma^2) &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} \frac{e^{-\frac{1}{2}Q''_{Pj}(\theta_j)(z-\theta_j)^2}}{\sqrt{2\pi}} dz \sqrt{Q''_{Pj}(\theta_j)} \\ &= \exp[-\tfrac{1}{2}\theta_j^2/(1 + (Q''_{Pj}(\theta_j))^{-1})] \frac{1}{\sqrt{1 + (Q''_{Pj}(\theta_j))^{-1}}} \end{aligned} \tag{5.10}$$

Specific adoption of sculpturing option (A), the pseudo-t, provides that

$$\theta_j = \sqrt{\frac{(n-1)^2}{(n-3/2)} \log [1 + (\frac{\phi(\theta_j)}{\sigma})^2 \frac{1}{n-2}]} \text{ sign } \phi(\theta_j) \tag{5.11}$$

and

solutions from (5.1). It has been discovered that care is required in the choice of the μ, σ^2 start when analyzing a few short data histories. Also, the straightforward integration is numerically ill-conditioned.

Method 2: Quadratic Approximation (Laplace's Method Approach).

An appealing approximation to the integral (5.1,b) is obtained by expressing of the $j^{\underline{th}}$ likelihood component as

$$L_j(\mu,\sigma^2;s_j,t_j) = \int_{-\infty}^{\infty} \frac{e^{-\frac{1}{2}z^2}}{\sqrt{2\pi}} e^{-Q_{Pj}(z)} dz \tag{5.4}$$

where, except for irrelevant parameter-free constants,

$$Q_{Pj}(z) = \lambda(z)t_j - s_j \ln\lambda(z) , \tag{5.5}$$

which has a qualitatively bowl-shaped appearance. Hence quadratically approximate $Q_P(z)$ as follows:

$$Q_{Pj}(z) \simeq Q_{Pj}(\theta_j) + \frac{1}{2}Q''_{Pj}(\theta_j)(z-\theta_j) , \tag{5.6}$$

θ_j being the solution of

$$Q'_P(z) = \lambda'(z)t_j - s_j \frac{\lambda'(z)}{\lambda(z)} = 0 , \tag{5.7}$$

so

$$\lambda(\theta_j) = s_j/t_j , \tag{5.8,a}$$

where

$$L_j(\mu,\sigma^2;s_j,t_j) \;=\; \int_{-\infty}^{\infty} \frac{e^{-\frac{1}{2}z^2}}{\sqrt{2\pi}}\, e^{-\lambda(z)t_j}\, \frac{(\lambda(z)t_j)^{s_j}}{s_j!}\, dz\ , \qquad (5.1,b)$$

and

$$\log \lambda(z) \;=\; \mu + \sigma\phi(z) \;=\; \mu + \sigma z\psi(z) \qquad (5.1,c)$$

Throughout what follows the sculpturing function, ψ or equivalently ϕ, is assumed given.

The integrals (5.1,b) cannot be carried out analytically. Owing to the appearance of $e^{-\frac{1}{2}z^2}$, the use of Gauss-Hermite numerical integration is suggested. In the notation of Abramowitz and Stegun (1968),

$$\sqrt{2\pi}\; L_j(\mu,\sigma^2;s_j,t_j) \;=\; \int_{-\infty}^{\infty} e^{-x^2}\, f_j(x)dx \;\simeq\; \sum_i w_i f_j(x_i)$$

where $z = \sqrt{2}x$ and

$$f(x_i) \;=\; e^{-\lambda(\sqrt{2}x_i)t_j}((\lambda(\sqrt{2}x_i)t_j)^{s_j}\,\frac{1}{s_j!}\ ; \qquad (5.2)$$

the x_i, w_i values are taken from tables. A grid search among μ, σ^2 values then reveals the approximate maximum likelihood

5. Superpopulation Parameter Estimation

In order to provide suitable pooled, shrunken, estimates of an individual item rate it is necessary to invoke estimates of super-population parameters. Unfortunately, the superpopulation variate values are not observed directly, but are contaminated by "Poisson noise;" this complicates the task of parameter estimation.

Natural approaches to the estimation problem are through moment matching, maximum likelihood or Bayesian approaches. We suggest variations on these themes that require different degrees of computer-intensive effort, and are of differing effectiveness. The methods advanced for consideration have been compared by simulations. The limited histories typical in various fields, e.g. in reliability and survival studies, and in nuclear plant risk, do not encourage faith in the validity of asymptotic error analyses without such corroboration.

5.1 Likelihood Estimation for the Log Sculptured-Normal Poisson (L/S-N/P) Model

Suppose that a time history of length t_j results in s_j events for item j ($j = 1,2,...,J$). The data is to be analyzed with reference to the general L/S-N/P model of (4.1), but for the present $\mu_j = \mu$, a constant; regression will be discussed later.

Method 1: Likelihood By Gauss-Hermite Integration.

The likelihood of the parameter μ and σ^2 given the data and the L/S-N/P model can be expressed as

$$L(\mu,\sigma^2;\underline{s},\underline{t}) = \prod_{j=1}^{J} L_j(\mu,\sigma^2,s_j,t_j) \qquad (5.1,a)$$

at this point, but treated as a tuning parameter: the smaller n ($n > 2$), or larger h ($h < 0.25$) the greater the effect of the weights upon discrepant observations. In practice, $n = 4$ has given satisfactory performance, as will be suggested by simulations and trial data analyses. For a similar analysis procedure in the more classical robustness context see the biweights for estimation of a distributional center, Mosteller and Tukey (1977), p. 353 ff.; our weight $\hat{w}_j(n)$ is essentially an influence function, with degree of observational influence adjusted by choice of n or h, corresponding to the parameter c in biweight technology.

parameter far from the appropriate superpopulation mean μ_j exerts a small influence on the shrinkage term, so the quoted estimate $\hat{\lambda}_j$ tends to be nearly the raw rate s_j/t_j. Interestingly, all of these extended tail models induce weights in excess of unity on observations with log-raw-rates close enough to μ_j; this might be called <u>over-shrinkage</u>, and is noticeable in Table 1, p. 79 of Berger (1984), wherein a normal likelihood is combined with a Cauchy (one d.f. t) prior. This very same effect has been pointed out by Tukey in (1974), p. 132.

An interpretable approximation to these log rate estimates is obtained by the following linearization: in (4.3) or (4.6) start with $\hat{\varepsilon}_j(1) = \ln(s_j/t_j)$ and turn the Newton-Raphson crank once, but evaluate $w_j(n)$ at $\hat{\varepsilon}_j(1)$, i.e. utilize (4.9). Then

$$\ln \hat{\lambda}_j(1) \equiv \hat{\varepsilon}_j(1) = \frac{(s_j)\ln(s_j/t_j)+(\mu_j/\sigma^2)\hat{w}_j(n)}{(s_j) + (1/\sigma^2)\hat{w}_j(n)}. \quad (4.10)$$

The term (s_j) is the delta-method estimate of $(\mathrm{Var}[\ln(s_j/t_j)])^{-1}$, so the estimate quoted is seen to be nearly a linear combination of the raw rate estimate and the individualized mean, with the shrinkage towards the latter influenced by the discrepancy $(\ln(s_j/t_j)-\mu_j)/\sigma$ as reflected in the weights $\hat{w}_j(n)$; small discrepancies tend to shrink the estimate towards the mean, while large discrepancies are tolerated, i.e. left largely without shrinkage so that the quoted estimate is nearly $\ln \hat{\lambda}_j \simeq \log(s_j/t_j)$. The parameter n or h in the superpopulation models is not estimated

$$\lambda_j \equiv e^{\mu+\sigma\phi(z_j)} = (s_j - z_i w_j(h)) \frac{1}{t_j} \tag{4.7}$$

and

$$w_j(h) = \frac{(1-4h)^{-3/4}}{(1+2hz_j^2)\, e^{hz_j^2}} \tag{4.8}$$

It can be seen that (4.3), (4.5), and (4.7) all can on occasion have two real solutions, corresponding to the possibility of two modes in the likelihood, or Bayesian posterior, for ϕ_j. Strict adherence to likelihood doctrine would force computation of each solution and a check to see which globally maximizes likelihood--a possible but tedious task. The same is true of a Laplace method approach, which requires modification to account for the bimodality. Consequently it is proposed to simply modify the estimate-dependent weights (4.4) and (4.6) to estimated weights that utilize the raw-data estimate $\ln(s_j/t_j)$ in place of $\hat{\phi}_j$, so in each case

$$\hat{w}_j(n) = \frac{C(n)}{1 + \left(\frac{\ln(s_j/t_j)-\mu_j}{\sigma}\right)^2 \left(\frac{1}{n-2}\right)} \tag{4.9}$$

which can be computed once; this modification leads to a unique, approximately Bayes, solution with reasonable properties. Of course both likelihood (or posterior) bimodality or the occurrence of a small weight value may suggest the need for model changes or other action.

Notice that in each case the weight $w_j(n)$ modifies the resulting estimate towards discrepancy tolerance: a value of the individualization

$$\ell_j(\phi_j;\mu_j,\sigma^2,s_j,t_j) = -\frac{(n-1)^2}{2(n-3/2)} \ln\left(1+\frac{\phi_j^2}{n-2}\right) - \lambda(\phi_j)t_j$$

$$+ s_j \ln\lambda(\phi_j), \tag{4.2}$$

and is expressed in terms of the pseudo t individualization $\phi_j \equiv \phi(z_j)$. Differentiation then yields the estimating equation for λ_j, or ϕ_j: if $\varepsilon_j = \mu_j + \sigma\phi_j$, then

$$\hat{\lambda}_j \equiv e^{\hat{\varepsilon}_j} = \left(s_j - \left(\frac{\mu_j-\hat{\varepsilon}_j}{\sigma^2}\right)w_j(n)\right)\frac{1}{t_j} \tag{4.3}$$

where

$$w_j(n) = \frac{(n-1)^2}{(n-3/2)(n-2)} \frac{1}{1+\left(\frac{\hat{\varepsilon}_j-\mu_j}{\sigma}\right)^2\frac{1}{n-2}} \tag{4.4}$$

(B): The log-likelihood is very similar to (4.2); one obtains only a slightly different weight:

$$\hat{\lambda}_j \equiv e^{\hat{\varepsilon}_j} = \left(s_j - \left(\frac{\mu_j-\hat{\varepsilon}_j}{\sigma^2}\right)w_j(n)\right)\frac{1}{t_j}, \tag{4.5}$$

with

$$w_j(n) = \left(\frac{n+1}{n-2}\right) \frac{1}{1+\left(\frac{\hat{\varepsilon}_j-\mu_j}{\sigma}\right)^2\frac{1}{n-2}} \tag{4.6}$$

(C): In this case the convenient individualization parameter is a normal deviate, z_j, where $\phi(z_j) = z_j \exp(hz_j^2)$:

(B) $\phi(z,n) = (\sqrt{(n-2)/n})\,t_n$, a true Student t with n deg. fr., again with unit variance if $n > 2$. Parenthetically, use of a finite varianc~ t is in no way essential.

(C) $\phi(z;h) = (1-4h)^{3/4} z e^{hz^2}$, $0 \leq h < \frac{1}{4}$ a member of Tukey's h-family; see Hoaglin (1983), and Tukey (1974).

There are other interesting and convenient such forms that give promise of providing multi-modal parametric representations; see Cobb (1983). All of our above forms yield distributions of $\varepsilon = \ln \tilde{\lambda}$ that are unimodal and symmetric, have variance σ^2, and are more stretch-tailed than the basic unit normal, $\tilde{z}$. As will be seen, this latter qualitative modification has a beneficial effect upon the rate estimator, reducing the tendency for indiscriminate shrinkage of apparently very discrepant observations.

4.1 Nearly Explicit Discrepancy-Tolerant (Controlled Shrinkage) Estimators of Rates.

The several sculptured normal representations just presented yield directly interpretable rate estimates by way of approximate likelihood maximization, as in (3.3). Results for options (A), (B), and (C) are sketched.

(A): Invert the pseudo-t to find

$$\exp(-z^2/2) = \left(1 + \frac{\phi^2}{n-2}\right)^{-(n-1)^2/2(n-3/2)}$$

Now the log-likelihood associated with observation j is proportional to

4. Discrepancy-Tolerant Versions of the Log-Normal (L/N/P) Model: The Log Sculptured-Normal Poisson (L/S-N/P).

It is natural to generalize the L/N/P model as follows. Again put

$$\underset{\sim}{\lambda}_j = \exp(\underset{\sim}{\varepsilon}_j) , \tag{4.1,a}$$

but now

$$\underset{\sim}{\varepsilon}_j = \ln \underset{\sim}{x}_j = \mu_j + \sigma\phi(\underset{\sim}{z}_j) \equiv \mu_j + \sigma \underset{\sim}{z}_j \psi(\underset{\sim}{z}_j) , \tag{4.1,b}$$

where $\mu_j = \mu + \underline{\beta}\,\underline{x}_j$, $\underset{\sim}{z}_j \sim N/(0,1)$, and $\psi(\underset{\sim}{z}_j) = \phi(\underset{\sim}{z}_j)/\underset{\sim}{z}_j$ is a <u>sculpturing function</u>. See Gaver (1983), and also Hoaglin (1983) for an account of certain such functions in the normal (Gaussian) context, attributable to Tukey, see (1974). Call $\phi(\underset{\sim}{z}_j)$ a <u>sculptured normal</u>. The purpose of (4.1,b) is to describe <u>stretch-tailed distributions</u> of log-rates, i.e. those that exhibit quite widely straggling exotic or extreme values, or outliers, both above and below the normal-like central part. Illustrations of well-behaved distributions of rates, as contrasted to straggling tailed distributions, and even multi-modal distributions, appear in Figs. 1, 2, and 3. Our methods are aimed at dealing with the effects of Fig. 2, and, to an extent, Fig. 3.

Here are some convenient sculpturings of the normal.

(A) $\phi(z;n) = \frac{z}{|z|}\sqrt{n-2}\left[\exp\left(z^2\frac{(n-3/2)}{(n-1)^2}\right)-1\right]^{1/2}$, a <u>pseudo-t</u>; an explicitly-invertible approximation to a true Student t with unit variance if $n > 2$. See Gaver and Kafadar (1984). This representation is used extensively later in the paper.

$$Q''(\hat{\theta}_j) = 1 + \sigma^2 s_j(\alpha\beta\hat{\theta}_j[\frac{1+(\hat{\phi}_j)^2/\alpha}{\hat{\phi}_j}]) , \qquad (5.34)$$

and now (5.26) can be evaluated for any μ, σ^2. The resulting approximate log likelihood, $\sum_{j=1}^{J} \log L_j(\mu,\sigma^2)$, with L_j as in (5.24) and above can be explored for maxima by numerical search. Alternatively, numerical integration of $\exp(-\ell(\mu,\sigma^2))$, see (5.26), with suitable (non-informative) priors for μ and σ^2 results in Bayes estimates; this ambitious numerical step has not yet been carried out.

A more detailed investigation may begin by precise determination of solutions to (5.29). Graphical analysis quickly reveals the possibility of three real solutions, two of which identify local maxima. It can be seen that, given μ and σ^2, the solution in terms of $\varepsilon = (\phi-\mu)/\sigma$ of the equation (4.5) modified with the weight (4.9) is a reasonable choice for $\hat{\phi}_j$, which is then converted to $\hat{\theta}_j$ by (4.3). Now

$$Q'(\hat{\theta}_j) = \hat{\theta}_j + (\lambda(\hat{\theta}_j)t - s_j)\phi'(\hat{\theta}_j) , \qquad (5.35)$$

and

$$Q''(\hat{\theta}_j) = 1 + \sigma^2 t_j(\phi'(\hat{\theta}_j))^2 + \sigma(\lambda(\theta_j)t_j - s_j)\hat{\phi}''(\hat{\theta}_j) \qquad (5.36)$$

where

$$\phi'(\hat{\theta}_j) = \alpha\beta \frac{\hat{\theta}_j}{\phi(\hat{\theta}_j)} (1 + \phi^2(\hat{\theta}_j)/\alpha) \qquad (5.37)$$

and

$$\phi''(\hat{\theta}_j) = \frac{1}{\phi(\hat{\theta}_j)}\{\alpha\beta(1+\phi^2(\hat{\theta}_j)/\alpha)(1+\alpha\beta\hat{\theta}_j^2) - (\phi'(\hat{\theta}_j))^2\} . \tag{5.38}$$

The Q_j-derivatives are introduced into (5.24), and the resulting expression is searched for maximizing values of μ and σ^2.

6. Simulation Testing

The procedures described for carrying out the estimation of superpopulation center (μ) and variance (σ^2), and the individualized, or selectively shrunken, estimates of rates, e.g. (4.3), have been appraised by simulations. In summary, the procedures described deliver satisfactory results; empirical distributions of estimates of μ and σ^2 appear to center close to the values input, and the average distance of selectively shrunken individualized rate estimates from the true (distance being mean squared error, median absolute deviation) often improve upon obvious competitors. Of course these statements apply only to the range of parameter values studied, which illustrate those encountered in certain nuclear power system probabilistic risk assessments. A brief selection of many simulations appears in the following tables and figures.

6.1 Simulation Design.

A simulation requires specification of the superpopulation form and parameters, the sample size, the exposure times t_j ($j = 1,2,\ldots,J$) and the sculpturing function $\phi(\cdot)$ used in rate production. Note that the latter function need not--and here will not--be the same as that used to construct individualized rate estimates.

Specification of the present simulation follows.

(a) Superpopulation form is a sculptured normal form (C), the Tukey h family, from which actual or "true" rates are easily constructed: $\lambda_{(j)} = \exp[\mu + \sigma\phi(z_{(j)})]$, where $z_{(j)}$ being the $j^{\underline{th}}$ ordered of J, i increasing magnitude) unit normal, where $\phi(z_{(j)}) = z_{(j)}\exp(hz^2_{(j)})$.

Simulation of a sample of J begins by obtaining J unit normal deviates, ordering them, and computing $\lambda_{(j)}$, $j = 1,2,\ldots,J$.

(b) Having $\lambda_{(j)}$, and having specified t_j, J realizations of independent Poisson random variables, or counts, are generated; s_j corresponds to mean $\lambda_{(j)} t_j$.

(c) The simulated observations (s_j, t_j; $j = 1,2,\ldots,J$) are analyzed according to the L/S-N/P model by Method 3 to obtain point estimates of superpopulation parameters μ and σ^2. The pseudo-t option (A) is used in the likelihood; parameter n has been treated as a tuner, either specified as low, e.g. n = 4, yielding highly restricted or selective shrinkage, or as high, e.g. n = 50, corresponding nearly to the more conventional log normal model. A numerical search procedure has been utilized to locate μ and σ^2 values.

(d) Individualized estimates $\hat{\lambda}_{(j)}$ are computed by these options, using estimated μ and σ^2 from (c):

$$\text{MLE:} \quad \hat{\lambda}_{(j)} = s_j/t_j \tag{6.1}$$

$$\text{SSP:} \quad \hat{\lambda}_{(j)} = \text{solution of (4.3) using } w_j(n) \equiv 1 \tag{6.2}$$

$$\text{RSP:} \quad \hat{\lambda}_{(j)} = \text{solution of (4.3) using the } \hat{w}_j(n) \text{ of (4.9);} \tag{6.3}$$

The abbreviations are, respectively, for maximum likelihood, simple shrinkage Poisson, and restricted shrinkage Poisson.

(e) Each case $(J, \mu, \sigma^2, h, (t_j), n)$ combination is independently simulated 200 times.

(e) Each case, i.e. (J,μ,σ^2,h,t_j,n) combination, is independently simulated 200 times for Table 6.1, and 100 times for Table 6.2. The mean-squared errors of estimators (6.1), (6.2), (6.3) are summarized--summary figures (sample mean, median, standard deviation, median absolute deviation, and mean squared error) for errors of estimate of the <u>smallest</u> true rate $\lambda_{(1)}$, the <u>median</u> true rate $\lambda_{(\frac{J+1}{2})}$, and the largest, $\lambda_{(J)}$: $\hat{\lambda}_{(j)} - \lambda_{(j)}$. Choice of these three values for true λ is enough to reveal the different behaviors of the candidate estimators: in general SSP and RSP both greatly improve upon simple MLE for centrist (median) λ values by borrowing strength, while RSP improves upon SSP at true λ extremes by refusing to overshrink.

Tables 6.1 and 6.2 summarize illustrative sets of simulation results. Note that the estimates of the superpopulation mean, μ, appear close to being unbiased, while those of the variance σ^2 appear to be consistently biased downwards in Table 6.1 ($J = 15$), and about right in Table 6.2 ($J = 45$). Standard error of estimate (square-roots of the variances of the empirical distributions of the corresponding parameters) are, not surprisingly, substantial; as is sensible, they decrease as J increases. Nevertheless, comparison of the simulated MSE figures for the various estimators suggest that RSP, especially for $n = 4$, has advantages: for the smallest rates, $\lambda_{(1)}$, and the largest, $\lambda_{(15)}$, RSP's mse more nearly resembles the MLE mse performance than does the more heavily shrunken SSP, particularly for $n = 50$ and 75 which imitate the action of a log-normal analysis; for middle values, $\lambda_{(8)}$ and $\lambda_{(23)}$, both RSP and SSP estimates shrink moderately, all improving substantially upon the MLE.

Table 6.1

Selected Mean Squared Error Comparisons
and Estimated Superpopulation Parameters

J = 15, h = 0.15, 200 Simulations

True Values	Estimated		Estimator	Mean Squared Error $\lambda_{(1)}$ (small)	$\lambda_{(8)}$ (median)	$\lambda_{(15)}$ (large
	(n = 4):	$\hat{\mu} = -0.97(0.41)$	RSP	0.016	0.019	0.33
		$\hat{\sigma}^2 = 0.17(0.15)$	SSP	0.020	0.020	0.34
$\mu = -1.0$						
$\sigma^2 = 0.25$			MLE	0.007	0.030	0.32
	(n = 50):	$\hat{\mu} = -0.98(0.45)$	RSP	0.019	0.020	0.35
		$\hat{\sigma}^2 = 0.18(0.15)$	SSP	0.019	0.020	0.35
	(n = 4):	$\hat{\mu} = -1.93(0.50)$	RSP	0.0050	0.0060	0.28
		$\hat{\sigma}^2 = 0.18(0.17)$	SSP	0.0060	0.0058	0.30
$\mu = -2.0$						
$\sigma^2 = 0.25$			MLE	0.0026	0.014	0.27
	(n = 50):	$\hat{\mu} = -1.93(0.52)$	RSP	0.0053	0.0057	0.30
		$\hat{\sigma}^2 = 0.20(0.18)$	SSP	0.0054	0.0057	0.30

Table 6.2

Selected Mean Squared Error Comparisons
and Estimated Superpopulation Parameters

$J = 45,\ h = 0.10,\ t_j = 5$; 100 Simulations

True Values	Estimated	Estimator	$\lambda_{(1)}$ (small)	$\lambda_{(23)}$ (median)	$\lambda_{(45)}$ (large)
	$(n=4): \hat{\mu} = 0.50 (0.25)$	RSP	0.030	0.067	2.65
	$\hat{\sigma}^2 = 0.41 (0.29)$	SSP	0.050	0.067	2.75
$\mu = 0.50$, $\sigma^2 = 0.35$		MLE	0.011	0.13	2.61
	$(n=75): \hat{\mu} = -0.56 (0.30)$	RSP	0.042	0.069	2.68
	$\hat{\sigma}^2 = 0.44 (0.28)$	SSP	0.044	0.069	2.71

7. Analysis of Data

Our methodology will now be applied to several sets of observational failure or event rate data. In each case estimated superpopulation parameters $\hat{\mu}$ and $\hat{\sigma}^2$ are quoted, as are the unpooled maximum likelihood individual rate estimates (MLE), the simple linearized shrunken (Bayes) estimates (SSP), and the discrepancy-tolerant restricted shrinkage estimates (RSP), along with the weights associated with each of the latter. It appears that the results so obtained contrast interestingly, with the RSP behaving in the discrepancy-tolerant manner anticipated, and with small weights influencing this behavior, especially in data sets for which J is substantial.

7.1 Ship System Failure Rates

The numbers of failures during one year experienced by each of J = 254 individual systems aboard a Navy ship have been furnished by Dr. R. Coile. It is provisionally assumed that all systems are exposed to failure throughout time, and that the failure process is nearly Poisson; neither assumption can be checked, but the analysis is of interest. The analytical results are in Table 7.1.

The heavy preponderance of zero and one-failure systems is recognized by both SSP and RSP, which nearly agree: both shrink in the same direction at this level. However, SSP continues to shrink towards low rate values even for the two units with relatively high observed rates, while RSP is far more discrepancy-tolerant, as dictated by the corresponding low weights. It is interesting that our procedures estimate, on the basis of a log-normal superpopulation, an expected failure rate of $\exp(\hat{\mu} + \frac{1}{2}\hat{\sigma}^2) = 0.54$, close to the pooled

TABLE 7.1

SHIP SYSTEM FAILURE RATES

$\hat{\mu} = -1.34, \quad \hat{\sigma}^2 = 1.46, \quad J = 254$

Number of Units	Failures	MLE	SSP (n = 75)	RSP (n = 4)	WT.
178	0	0.00	0.20	0.22	1.8
48	1	1.00	0.52	0.50	1.1
16	2	2.00	1.0	1.2	0.75
3	3	3.00	1.7	2.2	0.59
6	4	4.00	2.5	3.1	0.51
1	5	5.00	3.3	4.2	0.45
1	9	9.00	6.8	8.2	0.34
1	11	11.00	8.6	10.2	0.31

observed rate of 0.50. The median failure rate is $\exp(\hat{\mu}) = 0.26$, and since over one-half of the observations are zeros, it is also encouraging that the estimated median rate is near the estimated rate for systems with zero failures. Finally, a calculation of the probability of zero events for a gamma superpopulation moment-matched to the log-normal with the observed parameters yields 0.75, which may be compared to the observed fraction 178/254 = 0.70. Although the simple calculation is perhaps crude, the agreement is gratifying.

7.2 Loss of Feedwater Flow

Table 7.2 presents a set of data referring to the rates of loss of feedwater flow for a collection of nuclear power generation systems; see Kaplan (1983). The corresponding SSP, RSP derived rates are included. Once again, the units with very small (< 0.5) weights, namely Systems 1, 3, 7, 18, and 19, all display marked differences between RSP and SSP, with the resulting SSP estimates exhibiting shrinkage upwards far more extensive than those of the corresponding RSP Systems 11 and 23 have the highest observed rates, both have about the same times of exposures and nearly the same computed weights, and both RSP estimates are slightly less shrunken, thus closer to the MLE, than are those from SSP. The estimated median rate, calculated on the basis of a log-normal superpopulation, is $\exp(\hat{\mu}) = 2.56$; the fraction of MLE rates equal to or exceeding 2.6 is 16/30 = 0.53; the corresponding fractions of SSP and RSP rates is 15/30 = 0.5, in good agreement.

Table 7.2

Loss of Feedwater Flow Rates

$\hat{\mu} = 0.94 \quad \hat{\sigma}^2 = 0.31; \quad \hat{\lambda} = \exp(\hat{\mu} + \hat{\sigma}^2/2) = 2.99$

System	s(j)	t(j)	MLE	SSP (n = 75)	RSP (n = 4)	WT.
1	4	15	0.27	0.59	0.35	0.19
2	40	12	3.3	3.3	3.2	1.6
3	0	8	0.041	0.59	0.087	0.063
4	10	8	1.3	1.5	1.5	0.98
5	14	6	2.3	2.4	2.4	1.8
6	31	5	6.2	5.7	5.8	0.30
7	2	5	0.4	1.0	0.64	0.27
8	4	4	1.0	1.5	1.4	0.74
9	13	4	3.3	3.1	3.0	1.7
10	4	3	1.3	1.7	1.8	1.1
11	27	4	6.8	6.1	6.2	0.71
12	14	4	3.5	3.3	3.2	1.6
13	10	4	2.5	2.5	2.5	1.8
14	7	2	3.5	3.2	3.1	1.6
15	4	3	1.3	1.7	1.8	1.1
16	3	3	1.0	1.5	1.5	0.74
17	11	2	5.5	4.6	4.6	0.92
18	1	2	0.5	1.4	1.0	0.34
19	0	2	0.17	1.2	0.41	0.14
20	3	1	3.0	2.8	2.7	1.7
21	5	1	5.0	3.8	3.7	1.0
22	6	1	6.0	4.3	4.5	0.83
23	35	5	7.0	6.4	6.6	0.68
24	12	3	4.0	3.6	3.5	1.4
25	1	1	1.0	1.9	1.8	0.74
26	10	3	3.3	3.1	3.0	1.6
27	5	2	2.5	2.5	2.5	1.8
28	16	4	4.0	3.7	3.6	1.4
29	14	3	4.7	4.1	4.1	1.1
30	58	11	5.3	5.1	5.1	0.98

7.3 Globe Valve Leak Failures.

These data pertain to nuclear power plant globe valve leak-mode failures, categorized according to operator type, the source being NPRDS data; see Hill, et al (1984), p. 9, where the results of a gamma superpopulation analysis are presented and discussed. The data are categorized by operator type, and the between-category variability is described by a superpopulation. A more appropriate analysis would presumably be (known) category by category, with between-system variability within each known category described by a superpopulation. Table 7.3 describes the data and estimates, but also includes the gamma estimates of Hill et al.

Table 7.3

Globe Valve Leak Failure Rates

$\hat{\mu}(4) = 0.040$, $\hat{\sigma}^2(4) = 1.1$, $\hat{\mu}(75) = 0.18$, $\hat{\sigma}^2(75) = 1.08$

Category	s(j)	t(j)	MLE	Gamma PEB	SSP (n = 75)	RSP (n = 4)	WT.
1	31	236.9	0.131	0.134	0.138	0.136	0.60
2	157	115.9	1.35	1.35	1.35	1.35	1.7
3	30	36.8	0.815	0.823	0.816	0.825	1.8
4	13	7.60	1.71	1.67	1.63	1.62	1.6
5	7	5.47	1.28	1.27	1.22	1.23	1.8
6	7	1.69	4.14	3.35	3.38	3.51	0.96
7	0	1.12	0.00	0.411	0.47	0.55 (0.50)	1.0 (0.83)
8	0	0.55	0.00	0.559	0.59	0.78 (0.65)	1.6 (0.83)

The gamma-PEB and the present SSP-RSP methodologies give rather comparable results for the first six categories. The last two categories differ more strikingly, with the SSP-RSP procedure shrinking somewhat more extensively than the gamma, the RSP weights, especially that for category 8, are surprisingly high; this is believed to be the result of the necessity of approximating the assessment of discrepancy on the log-scale for $s_7 = s_8 = 0$. If the experiences for categories 7 and 8 are pooled in order to compute weights, then the perhaps more acceptable numbers in parentheses result. Although the weights placed on the apparently discrepant rates for categories 1, 7, and 8 are not as striking as might be wished, they are of interest. It must be recognized that $J = 8$ is a very small group or "sample."

8. Acknowledgements

The author is grateful to J. Mayberry for helpful editorial comments, and also to L. Uribe for programming assistance.

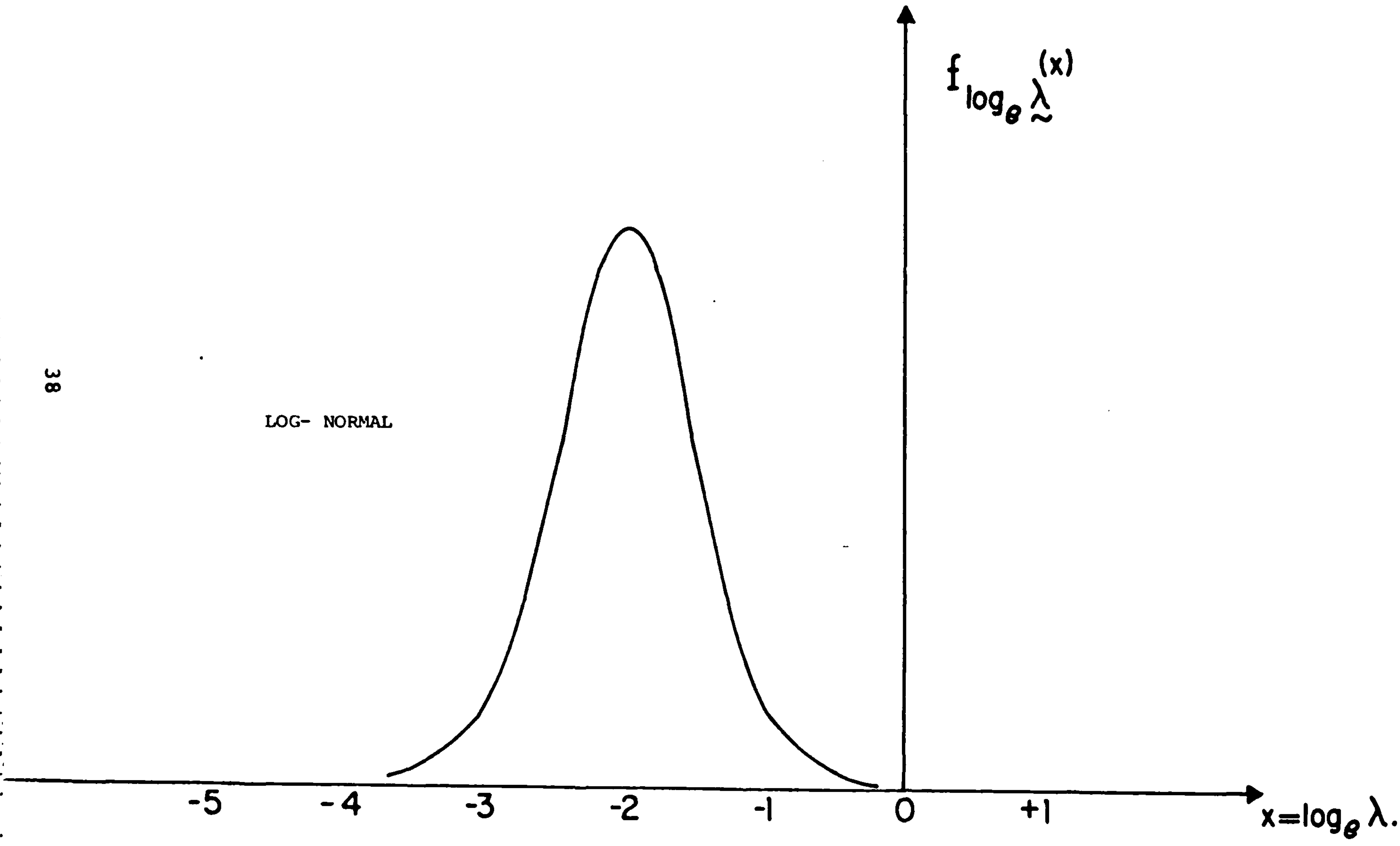

$f_{\log_e \underset{\sim}{\lambda}}(x)$
LOG- NORMAL
-5
-4
-3
-2
-1
0
+1
$x = \log_e \lambda.$

FIGURE 1

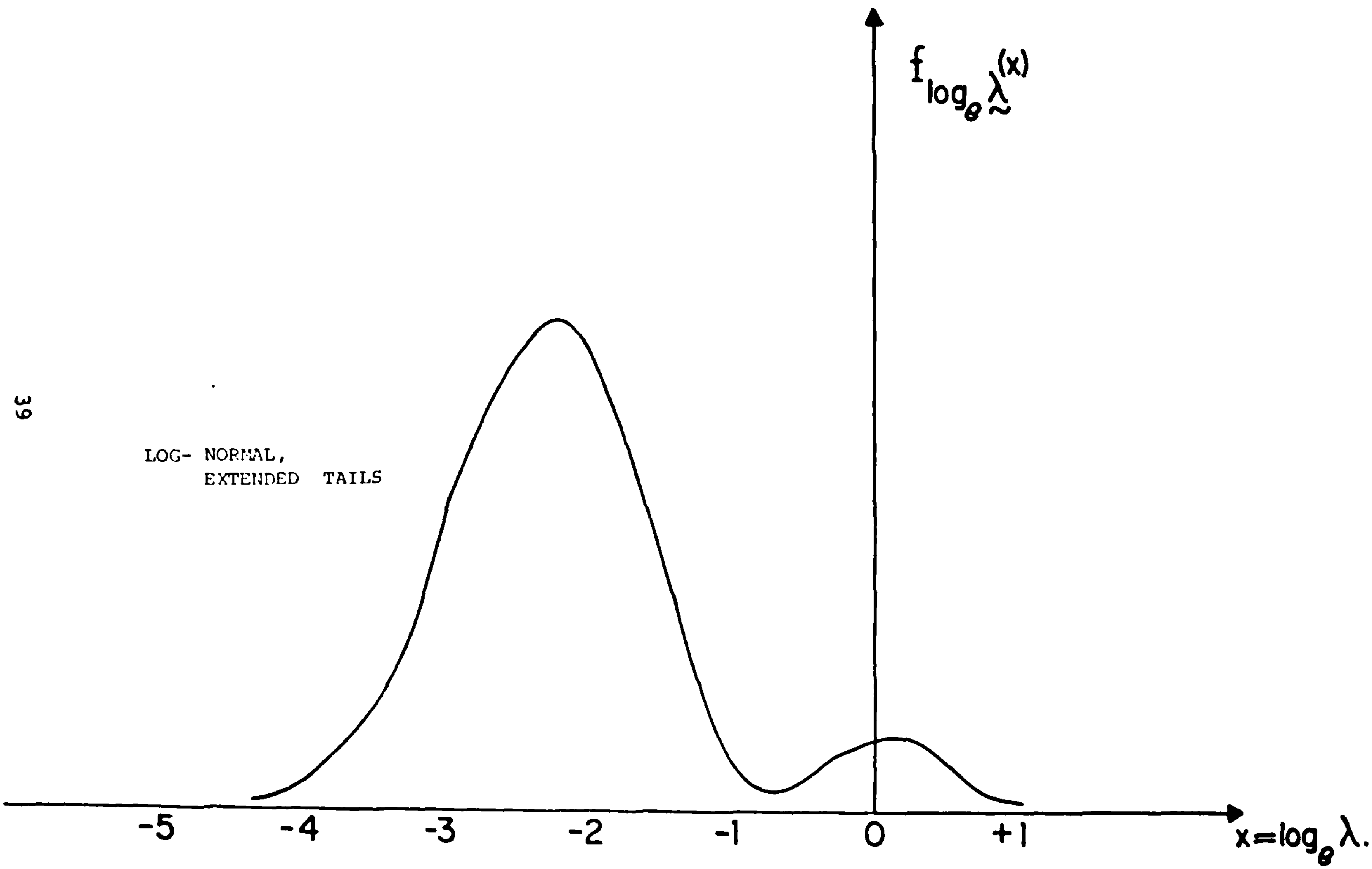

$f_{\log_e \underset{\sim}{\lambda}}(x)$
LOG- NORMAL,
EXTENDED TAILS
-5
-4
-3
-2
-1
0
+1
$x = \log_e \lambda.$

FIGURE 2

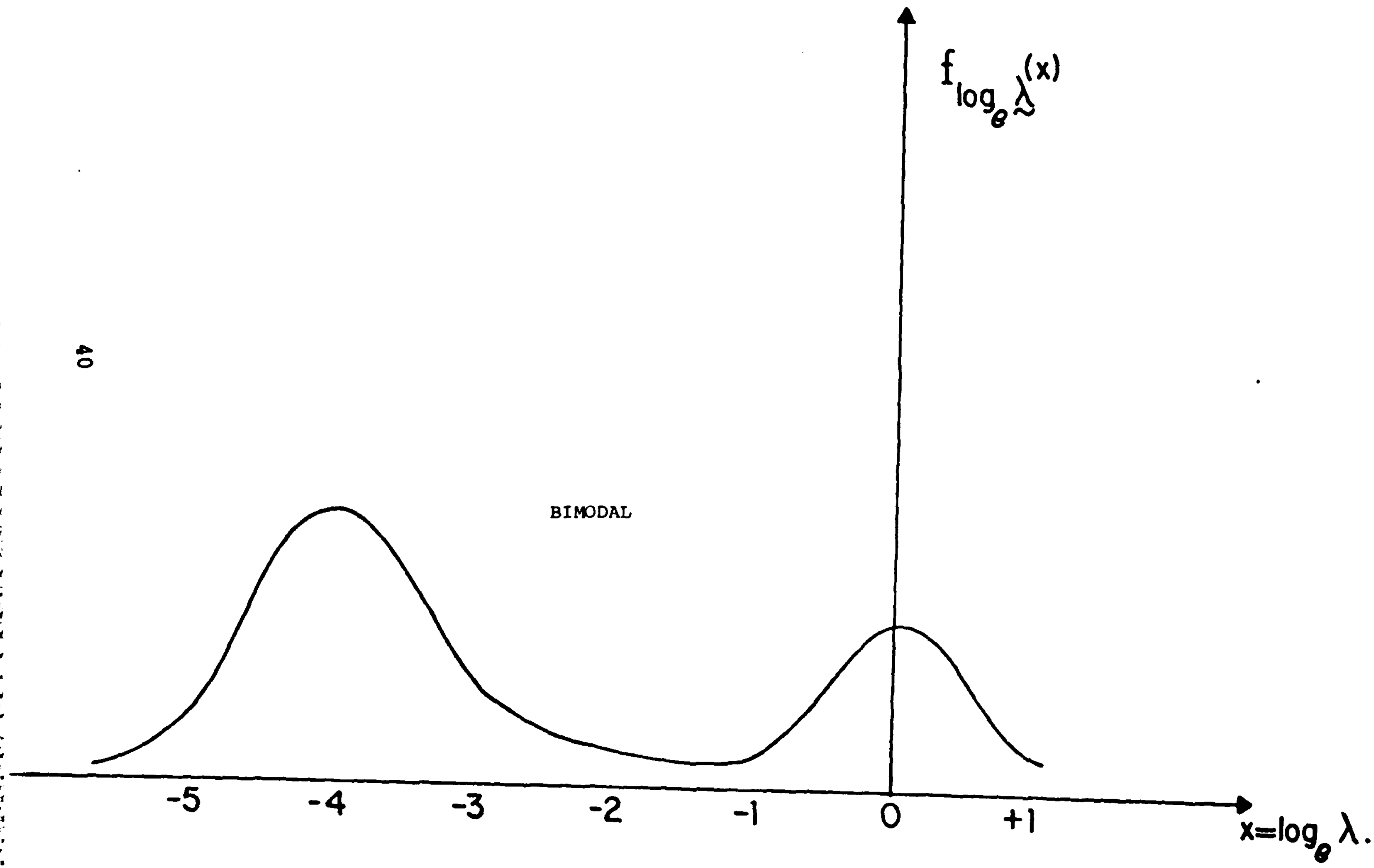


FIGURE 3

REFERENCES

Abramowitz, M. and Stegun, I.A. (1968). Handbook of Mathematical Functions. National Bureau or Standards, AMS SS. Washington, D.C.

Albert, J.H. (1979). "Robust Bayes Estimation." Purdue University Dept. of Statistics Mimeo. Series #79-9.

Belsley, D.A., Kuh, E. and Welsch, R.E. (1980). Regression Diagnostics. New York: John Wiley.

Berger, J.). (1980). Statistical Decision Theory. Springer-Verlag.

Berger, J.O. (1984). "The robust Bayesian viewpoint;" in Robustness of Bayesian Analyses, J.B. Kadane (ed.), Elsevier.

Box, G.E.P. (1980). "Sampling and Bayes' inference in scientific modelling and robustness." Journal of the Royal Statistical Soc., 143, A, pp. 383-430.

Burridge, J. (1981). "Empirical Bayes analysis of survival time data." J.R. Statist. Soc. B, 43, No. 1, pp. 65-76.

Cobb, L., Koppstein, P., and N.H. Chen (1983). "Estimation and moment recursion relations for multimodal distributions of the exponential family." Journal of the American Statistical Association, 78, pp. 124-130.

Copas, J.B. (1984). "A Computer-Intensive Non-Parametric Estimate of a Mixing Distribution, with Application to Poisson Failure Data." Monterey: Naval Postgraduate School. Technical Report.

Cox, D.R., and Isham, V. (1980). Point Processes. Chapman and Hall. London.

Cox, D.R., and Lewis, P.A.W. (1966). The Statistical Analysis of Series of Events. London: Methuen.

Deely, J.J. and Lindley, D.V. (1981). "Bayes Empirical Bayes." Journal of the American Statistical Association. Vol. 1, 76, No. 376, pp. 833-841.

Efrcn, B. and Morris, C. (1973). "Stein's Estimation Rule and Its Competitors--An Empirical Bayes Approach." Journal of the American Statistical Association. Vol. 68, No. 341, pp. 117-130.

Gaver, D.P. (1963). "Random hazard in reliability problems." Technometrics, 5, pp. 211-226.

Gaver, D.P., and Kafadar, K. (1984). "A retrievable recipe for inverse t." American Statistician, pp. 308-311.

Hill, J.R., Heger, A.S. and Koen, B.V. (1984). "The Application of Stein and Related Parametric Empirical Bayes Estimators to the Nuclear Plant Reliability Data System." Austin, Texas: The Univ. of Texas. NUREG/CR-3637.

Hoaglin, G.C. (1983). "g and h distributions." Encyclopedia of Statistics.

Kaplan, S. (1983). "On a two-stage Bayesian procedure for determining failure rates from experiential data." IEEE Transactions on Power Apparatus and Systems, Vol. PAS-102, No. 1, pp. 195-202.

Lehoczky, J.P. (1984). "Random parameter stochastic process models of criminal careers." Carnegie-Mellon Univ., Pittsburgh, Pa.; Dept. of Statistics Technical Report No. 343.

Martz, H.F., Beckman, R.J. and McInteer, C.R. (1982). "FRAC (Failure Rate Analysis Code): A Computer Program for Analysis of Variance of Failure Rates." Los Alamos, N.M.: Los Alamos National Laboratory. NUREG/CR-2434, LA-9116-MS.

Morris, C. (1982). "Natural exponential families with quadratic variance functions: statistical theory." Annals of Statistics, Vol. 11, No. 2, pp. 515-529.

Morris, C. (1983). "Parametric empirical Bayes inference: theory and applications." Journal of the American Statistical Association. Vol. 78, No. 381 (with discussion), pp. 47-65.

Mosteller, F., and Tukey, J.W. (1977). Data Analysis and Regression. Addison-Wesley: Reading, Mass.

Mosteller, F. and Wallace, D.L. (1964). Inference and Disputed Authorship: The Federalist. Addison Wesley: Reading, Mass.

Rasmussen, N. et al (1975). "Reactor Safety Study: An Assessment of Accident Risks in U.S. Commercial Nuclear Power Plants." NUREG-75/014, WASH 1400.

Reynolds, D.S., and Savage, I.R. (1971). "Random wear models in reliability theory." Adv. Appl. Prob., 3, pp. 229-248.

Robbins, H. (1983). "Some thoughts on empirical Bayes estimation." Annals of Statistics. Vol. 11, No. 3, pp. 713-723.

Tierney, L. and Kadane, J.B. (1984). "Accurate approximations for posterior moments and marginals." Tech. Report, Dept. of Statistics, Univ. of Minn. and Dept. of Statistics and Social Science, Carnegie-Mellon, Univ. Tech. Report.

Tukey, J.W. (1974). "Named and faceless values: an initial exploration in memory of Prasanta C. Mahalanobis." Sankhya, Series A, 36, pp. 125-176.

	NO. OF COPIES
Dr. D. F. Daley Statistics Dept. (I.A.S) Australian National University Canberra A.C.T. 2606 AUSTRALIA	1
Prof. F. J. Anscombe Department of Statistics Yale University, Box 2179 New Haven, CT 06520	1
Dr. David Brillinger Statistics Department University of California Berkeley, CA 94720	1
Dr. R. Gnanadesikan Bell Core 435 South Street Morris Township, NJ 07960	1
Prof. Bernard Harris Dept. of Statistics University of Wisconsin 610 Walnut Street Madison, WI 53706	1
Dr. D. R. Cox Department of Mathematics Imperial College London SW7 ENGLAND	1
Dr. A. J. Lawrance Dept. of Mathematical Statistics University of Birmingham P. O. Box 363 Birmingham B15 2TT ENGLAND	1
Professor W. M. Hinich University of Texas Austin, TX 78712	1
Dr. John Copas Dept. of Mathematical Statistics University of Birmingham P. O. Box 363 Birmingham B15 2TT ENGLAND	1

P. Heidelberger 1
IBM Research Laboratory
Yorktown Heights
New York, NY 10598

Prof. M. Leadbetter 1
Department of Statistics
University of North Carolina
Chapel Hill, NC 27514

Dr. D. L. Iglehart 1
Dept. of Operations Research
Stanford University
Stanford, CA 94350

Dr. D. Vere-Jones 1
Dept. of Mathematics
Victoria University of Wellington
P. O. Box 196
Wellington
NEW ZEALAND

Prof. J. B. Kadane 1
Dept. of Statistics
Carnegie-Mellon
Pittsburgh, PA 15213

Prof. J. Lehoczky 1
Department of Statistics
Carnegie-Mellon University
Pittsburgh, PA 15213

Dr. J. Maar (R51) 1
National Security Agency
Fort Meade, MD 20755

Dr. M. Mazumdar 1
Dept. of Industrial Engineering
University of Pittsburgh
Pittsburgh, PA 15235

Prof. M. Rosenblatt 1
Department of Mathematics
University of California-San Diego
La Jolla, CA 92093

Prof. Rupert G. Miller, Jr. 1
Statistics Department
Sequoia Hall
Stanford University
Stanford, CA 94305

Prof. I. R. Savage 1
Dept. of Statistics
Yale University
New Haven, CT 06520

Dr. Paul Shaman 1
National Science Foundation
Mathematical Sciences Section
1800 G. Street, NW
Washington, D. C. 20550

Prof. W. R. Schucany 1
Dept. of Statistics
Southern Methodist University
Dallas, TX 75222

Prof. D. C. Siegmund 1
Dept. of Statistics
Sequoia Hall
Stanford University
Stanford, CA 94305

Prof. H. Solomon 1
Department of Statistics
Sequoia Hall
Stanford University
Stanford, CA 94305

Dr. Ed Wegman 1
Statistics & Probability Program
Code 411(SP)
Office of Naval Research
Arlington, VA 22217

Dr. P. Welch 1
IBM Research Laboratory
Yorktown Heights, NY 10598

Dr. Marvin Moss 1
Office of Naval Research
Arlington, VA 22217

Dr. Roy Welsch 1
Sloan School
M. I. T.
Cambridge, MA 02139

Pat Welsh 1
Head, Polar Oceanography Branch
Code 332
Naval Ocean Research & Dev. Activity
NSTL Station, MS 39529

Dr. Douglas de Priest 1
Statistics & Probability Program
Code 411(SP)
Office of Naval Research
Arlington, VA 22217

Dr. Morris DeGroot 1
Statistics Department
Carnegie-Mellon University
Pittsburgh, PA 15235

Prof. J. R. Thompson 1
Dept. of Mathematical Science
Rice University
Houston, TX 77001

Prof. J. W. Tukey 1
Statistics Department
Princeton University
Princeton, NJ 08540

Dr. Daniel H. Wagner 1
Station Square One
Paoli, PA 19301

Dr. Colin Mallows 1
Bell Telephone Laboratories
Murray Hill, NJ 07974

Dr. D. Pregibon 1
Bell Telephone Laboratories - AT&T
Murray Hill, NJ 07974

Dr. Jon Kettenring 1
Bell Core
435 South Street
Morris Township, NJ 07960

Dr. David L. Wallace 1
Statistics Dept.
University of Chicago
5734 S. University Ave.
Chicago, IL 60637

Dr. F. Mosteller 1
Dept. of Statistics
Harvard University
Cambridge, MA 02138

Dr. S. R. Dalal 1
Bell Laboratories - AT&T
Mountain Avenue
Murray Hill, NJ 07974

Prof. Donald P. Gaver 20
Code 55Gv
Naval Postgraduate School
Monterey, California 93943-5100

Prof. Patricia Jacobs 1
Code 55Jc
Naval Postgraduate School
Monterey, California 93943-5100

Dr. Guy Fayolle 1
I.N.R.I.A.
Dom de Voluceau-Rocquencourt
78150 Le Chesnay Cedex
FRANCE

Dr. M. J. Fischer 1
Defense Communications Agency
1860 Wiehle Avenue
Reston, VA 22070

Prof. George S. Fishman 1
Curr. in OR & Systems Analysis
University of North Carolina
Chapel Hill, NC 20742

Prof. Guy Latouche 1
University Libre Bruxelles
C. P. 212
Blvd De Triomphe
B-1050 Bruxelles
BELGIUM

Library 4
Code 1424
Naval Postgraduate School
Monterey, CA 93943-5100

Dr. Alan F. Petty 1
Code 7930
Naval Research Laboratory
Washington, DC 20375

Prof. Bradley Efron 1
Statistics Department
Sequoia Hall
Stanford University
Stanford, CA 94305

Prof. Carl N. Morris 1
Dept. of Mathematics
University of Texas
Austin, TX 78712

Dr. John E. Rolph 1
RAND Corporation
1700 Main Street
Santa Monica, CA 90406

Prof. Linda V. Green 1
Graduate School of Business
Columbia University
New York, NY 10027

Dr. David Burman 1
Bell Laboratories - AT&T
Mountain Avenue
Murray Hill, NJ 07974

Dr. Ed Coffman 1
Bell Laboratories - AT&T
Mountain Avenue
Murray Hill, NJ 07974

Prof. William Jewell 1
Operations Research Department
University of California, Berkeley
Berkeley, CA 94720

Dr. Tom A. Louis 1
Biostatics Department
Harvard School of Public Health
677 Huntington Avenue
Boston, MA 02115

Dr. Nan Laird 1
Biostatics Department
Harvard School of Public Health
677 Huntington Avenue
Boston, MA 02115

Dr. Marvin Zelen 1
Biostatics Department
Harvard School of Public Health
677 Huntington Avenue
Boston, MA 02115

Dr. John Orav 1
Biostatics Department
Harvard School of Public Health
677 Huntington Avenue
Boston, MA 02115

Prof. R. Douglas Martin 1
Chairman
Department of Statistics, GN-22
University of Washington
Seattle, WA 98195

Prof. W. Stuetzle 1
Department of Statistics
University of Washington
Seattle, WA 98195

Prof. F. W. Mosteller 1
Department of Statistics
Harvard University
1 Oxford Street
Cambridge, MA 02138

Dr. D. C. Hoaglin 1
Department of Statistics
Harvard University
1 Oxford Street
Cambridge, MA 02138

Prof. N. D. Singpurwalla 1
George Washington University
Washington, D. C. 20052

Center for Naval Analyses 1
2000 Beauregard Street
Alexandria, VA 22311

Prof. H. Chernoff 1
Department of Mathematics
M. I. T.
Cambridge, MA 02139

Dr. T. J. Ott 1
Bell Core
435 South Street
Morris Township, NJ 07960

END

FILMED

11-85

DTIC